

AD \_\_\_\_\_

Award Number: DAMD17-01-1-0390

TITLE: Identification of Potential Biomarkers for the Early  
Diagnosis of Breast Cancer

PRINCIPAL INVESTIGATOR: Peiqing Sun, Ph.D.

CONTRACTING ORGANIZATION: The Scripps Research Institute  
La Jolla, CA 92037

REPORT DATE: September 2004

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command  
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;  
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

20050204 145

# REPORT DOCUMENTATION PAGE

Form Approved  
OMB No. 074-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

|   |   |  |  |  |
|---|---|--|--|--|
| <b>1. AGENCY USE ONLY</b><br>(Leave blank)  |   | <b>2. REPORT DATE</b><br>September 2004                        | <b>3. REPORT TYPE AND DATES COVERED</b><br>Annual (1 Sep 2003 - 31 Aug 2004) |  |
| <b>4. TITLE AND SUBTITLE</b><br>Identification of Potential Biomarkers for the Early Diagnosis of Breast Cancer   |   |  | <b>5. FUNDING NUMBERS</b><br>DAMD17-01-1-0390                                |  |
| <b>6. AUTHOR(S)</b><br><br>Peiqing Sun, Ph.D.   |   |  |  |  |
| <b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b><br>The Scripps Research Institute<br>La Jolla, CA 92037<br><br>E-Mail: pgsun@scripps.edu  |   |  | <b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>                              |  |
| <b>9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b><br>U.S. Army Medical Research and Materiel Command<br>Fort Detrick, Maryland 21702-5012  |   |  | <b>10. SPONSORING / MONITORING AGENCY REPORT NUMBER</b>                      |  |
| <b>11. SUPPLEMENTARY NOTES</b>  |   |  |  |  |
| <b>12a. DISTRIBUTION / AVAILABILITY STATEMENT</b><br>Approved for Public Release; Distribution Unlimited  |   |  | <b>12b. DISTRIBUTION CODE</b>  |  |
| <b>13. ABSTRACT (Maximum 200 Words)</b><br><br>The successful treatment of breast cancer requires detection of the disease at early stages. In this application, we propose to systematically isolate secreted and cell surface proteins (trafficked proteins) with increased expression in early stage breast tumors. These proteins are candidates for blood-borne markers and cell surface markers that can be used in the routine screening of early stage breast cancer. We proposed to isolate all trafficked proteins from breast tumors using a functional approach. Those with increased expression in tumors will be identified by microarray analyzed. A cDNA library has been constructed using human breast tumor samples, and the secretion trap screen has been performed. We have obtained a human breast tumor library in which the trafficked proteins have been enriched. Based on the sequencing analysis of a limited number of clones, we conclude that approximately 68% of the library is represented by trafficked proteins. Trafficked proteins found in this library include receptors, ion channels, transporters, proteases, cytokines, chemokines, adhesion molecules and carrier proteins, some of which are known to be overexpressed in human breast cancer. Micro-array analysis will be carried out to identify the trafficked proteins with increased expression in human breast tumor tissues. |   |  |  |  |
| <b>14. SUBJECT TERMS</b><br><br>Mammalian Genetics; Retrovirus; Trafficked Proteins; Secretion  |   |  | <b>15. NUMBER OF PAGES</b><br>14   |  |
|   |   |  | <b>16. PRICE CODE</b>  |  |
| <b>17. SECURITY CLASSIFICATION OF REPORT</b><br>Unclassified  | <b>18. SECURITY CLASSIFICATION OF THIS PAGE</b><br>Unclassified | <b>19. SECURITY CLASSIFICATION OF ABSTRACT</b><br>Unclassified | <b>20. LIMITATION OF ABSTRACT</b><br>Unlimited                               |  |

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)  
Prescribed by ANSI Std. Z39-18  
298-102

Identification of Potential Biomarkers for the  
Early Diagnosis of Breast Cancer  
Peiqing Sun, Ph.D.

**Table of Contents**

|  |               |
|--|---------------|
| <b>Cover.....</b>                        | <b>Cover</b>  |
| <b>SF 298.....</b>                       | <b>SF 298</b> |
| <b>Introduction.....</b>                 | <b>1</b>      |
| <b>Body.....</b>                         | <b>1</b>      |
| <b>Key Research Accomplishments.....</b> | <b>8</b>      |
| <b>Reportable Outcomes.....</b>          | <b>8</b>      |
| <b>Conclusions.....</b>                  | <b>8</b>      |
| <b>References.....</b>                   | <b>9</b>      |
| <b>Appendices.....</b>                   | <b>10</b>     |

## Introduction

The successful treatment of breast cancer requires detection of the disease at early stages. Currently the diagnostic tools for breast cancer suffer from drawbacks. For example, mammography can miss small lesions, and sometimes can induce tumors in certain patients. Tests of biopsied tumor samples can only be performed in symptomatic patients when the tumor has been identified. Certain blood-borne markers (such as PSA) have been used in affordable, routine serum tests to diagnose cancer in asymptomatic patients. However, these markers, when used individually, can detect the presence of only a certain percentage, but not all, of a particular type of tumors. This argues that a panel of blood-borne markers should be identified and used for the diagnosis of each type of cancer. It is conceivable that multiple blood-borne markers, when used in combination, could predict the existence of tumors more accurately and more successfully. Once the presence of a tumor is predicted, cell surface markers can be used in imaging analysis to identify the location of the lesion.

In this application, we propose to systematically isolate secreted and cell surface proteins (trafficked proteins) with increased expression in early stage breast tumors. These proteins are candidates for blood-borne markers and cell surface markers that can be used in the routine screening of early stage breast cancer. We will first isolate all secreted and cell surface proteins from breast tumors of multiple patients, using a functional approach we have designed and validated. Next, the expression levels of these proteins in normal and early stage breast tumor tissues will be compared, and those with increased expression in tumors will be identified and analyzed. These studies will generate a pool of potential biomarkers, which can be evaluated through further studies for their use in the early diagnosis of breast cancer.

## Body

### *1. Construction of a deep, representational secretion trap library from multiple breast tumors.*

In the original proposal, we planned to construct a secretion trap library from 40 breast tumor samples. However, due to the availability of tumor samples from the Scripps Clinic, we were only able to collect 20 tumors. A random-primed cDNA library was constructed in the pTRAPI vector from these tumor samples. Initially, total RNA was isolated from these 20 tumors, and the integrity of RNA was examined by Northern blot analysis using a cDNA fragment encoding GAPDH gene as probe. Five RNA samples were found degraded, and therefore was discarded. Equal amount of each of the remaining 15 samples was combined, and mRNA was subsequently purified by affinity chromatography with oligo-(dT) cellulose resin. Using Northern blot analysis, we confirmed that the signal representing GAPDH mRNA had been enriched by more than 10 folds in purified mRNA, as compared to the original total RNA sample. cDNA was then synthesized using the Stratagene  $\lambda$ -ZAP cDNA synthesis kit. Instead of the oligo-dT primer provided with the kit, we used during cDNA synthesis a random primer containing 6 non-hydrolyzable, thiolated nucleotides. These thiolated nucleotides

effectively protected the XhoI restriction site present in the primer from being degraded by DNA polymerase I that was used in the second strand cDNA synthesis. After synthesis, cDNA was size-fractionated. cDNA species of 300-2,000 base pairs (bp) in length was selected, and ligated directionally into the EcoRI and XhoI sites of the pTRAPI vector, upstream of the CD8 reporter gene. The ligated library DNA was transformed into DH10B bacterial cells by electroporation, amplified in bacteria and purified in large quantity.

Based on the number of bacterial colonies obtained after the initial transformation of ligated library DNA, we estimated that the primary library contained  $1.5 \times 10^7$  independent clones. To confirm the quality of the library, 20 clones were randomly

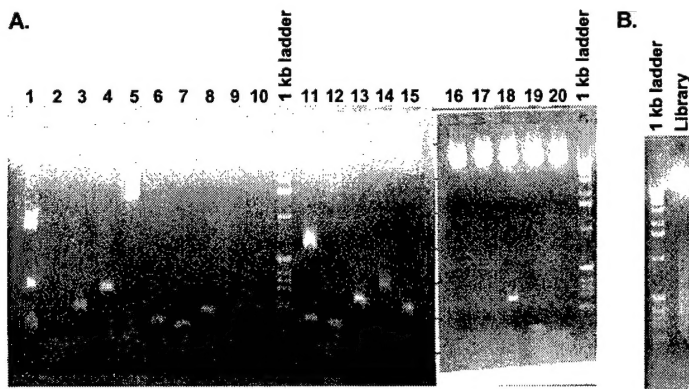


Fig. 1 Restriction analysis of the pTRAP-breast tumor library. A, Plasmid DNA was isolated from 20 randomly selected clones in the library, digested with EcoRI and XhoI to release the cDNA inserts, and analyzed by electrophoresis on 1% agarose gel. Based on the inserts from these 20 clones, it was estimated that 75% (15 out of 20) of the clones within this library contained cDNA inserts with an average size of 430 bp. B., 1  $\mu$ g of the total pTRAP-breast tumor library DNA was digested with EcoRI and XhoI, and analyzed by electrophoresis on 1% agarose gel.

selected from the primary library and evaluated. Restriction analysis using EcoRI and XhoI revealed that 75% (15) of these 20 clones contained cDNA inserts (Fig. 1A). From the sizes of cDNA inserts present in these clones, it was estimated that the average size of cDNA inserts in the library was approximately 430 bp. When digested with EcoRI and XhoI, cDNA

inserts released from the total library DNA migrated as a smear on agarose gel, ranging from 100 bp to 2,000 bp in size (Fig. 1B). The lack of discrete bands among the cDNA species suggested that cDNA of different sizes was represented equally in this library. Based on these results, we concluded that the human breast tumor library we had constructed was of good quality and was suitable for the subsequent secretion trap screen.

## 2. Isolation of cDNA encoding trafficked proteins from the breast tumor library using the secretion trapping screen.

We constructed a secretion trap library with  $1.5 \times 10^7$  independent clones from 15 human breast tumor samples. Two secretion trap screens were performed using this library to isolate cDNA fragments encoding the signal sequences of the trafficked proteins in human breast tumor tissues.

### Screen #1.

In our first screen, we packaged the library as a whole into retroviruses by transfecting the library DNA into a retroviral packaging cell line, LinX-A. 100 10-cm plates of 293 cells, each containing  $2 \times 10^6$  cells, were infected with the library viruses. A parallel infection with a control virus expressing the LacZ gene indicated an infection efficiency of 20-30%. Therefore, we estimated that at least  $4 \times 10^7$  cells had been infected, which was sufficient to cover the complexity of the library with a more than 2-fold redundancy. After purification of infected cells by hygromycin treatment, cells were subjected to selection with magnetic beads coupled to an anti-CD8 antibody. Because the yield of the first round of selection was expected to be low due to the low percentage of CD8 positive cells in the initial population, we decided to perform the selection in two large pools in order to shorten the period of recovery before the cells were ready for the second round.  $10^7$  cells were detached from each plate with 3 mM EDTA/EGTA and combined into 2 pools, with each pool containing  $5 \times 10^8$  cells from 50 plates. Each pool was selected with 2.5 ml of Dynabead M-450 CD8 in 50 ml of phosphate-buffered saline (PBS) containing 5% fetal calf serum and 0.6% sodium citrate in 50 ml Falcon tubes. After washing with the same buffer, the bound cells were eluted with 500  $\mu$ l of DETACHaBEAD, and plated into four 10-cm plates. The subsequent rounds of selection was performed in Eppendorf tubes with 50  $\mu$ l of Dynabead M-450 CD8 in a total volume of 1 ml, and the bound cells were eluted with 10  $\mu$ l of DETACHaBEAD. After each selection cycle, a portion of the recovered cells was immuno-stained with an anti-CD8 antibody coupled to a fluorescent dye (FITC), and the percentage of CD8 positive cells was determined by flow cytometry. After four rounds of selection with CD8 beads, the CD8 positive cells were enriched to about 80% for both pools.

The integrated proviruses were excised by treating the genomic DNA isolated from selected cells with Cre recombinase, and recovered after electroporation into *E. coli*. 20 single colonies were picked, and the plasmid DNA was isolated and sequenced to determine the identity of the cDNA fragment in each clone. Unfortunately, we found that both pools were dominated by 2-3 cDNA species. This finding indicated that this secretion trap screen failed due to the overgrowth of a few founder cells that had acquired growth advantage. Thus, the lesson we have learned from our first screen is that cells transduced with the library need to be kept as separate as possible during selection and growth.

#### Screen #2.

The strategy for our second secretion trap screen was designed to keep the cells transduced with the library in as many pools as possible during the selection process. In addition, we divided the library into pools of  $2 \times 10^5$  before packaged into viruses, in order to reduce the possibility that the growth-promoting cDNAs are present in all pools of the cells transduced with the library. The glycerol stock of the library was divided into 96 pools, each containing  $2 \times 10^5$  bacterial cells. Plasmids were isolated from these pools using a Qiagen robot. These pools of DNA were transfected separately into LinX-A packaging cells, and the resulting virus pools were used to infect 293 cells in 96 10-cm plates, each containing  $2 \times 10^6$  cells. At an infection efficiency of 20-30% (as determined in a parallel infection with a control virus expressing the LacZ gene), we estimated that

each pool of the library had been transduced into  $4 \times 10^5$  cells, and that the total number of cells transduced by the library was  $3.84 \times 10^7$ . After purification of transduced cells with hygromycin,  $10^7$  of cells from each pool were selected separately with 50  $\mu$ l of anti-CD8 beads in a total volume of 1 ml in an eppendorf tube, and the bound cells were eluted with 10  $\mu$ l of DETACHaBEAD. Eluted cells from each set of 4 plates were combined and plated into a fresh 10-cm plate for recovery. Therefore, after the first round of selection, the number of pools was reduced to 24, which were renamed from A to X. These 24 pools were kept separate during the following rounds of selection, which again was performed in Eppendorf tubes with 50  $\mu$ l of Dynabead M-450 CD8 in a total volume of 1 ml.

After each round of selection, a portion of the recovered cells was stained with an anti-CD8 antibody coupled to FITC, and the percentage of CD8 positive cells were determined by flow cytometry. After 4 rounds of selection, most of the pools contained more than 40% of CD8 positive cells, with the exception of pools S and V (Table 1 and Fig. 1A). Three pools, H, K and O, were contaminated at various stages during selection, and were excluded from further analysis. As controls, we also performed parallel selection with 293 cells transduced with either pTRAP alone or a combination of pTRAP containing a CD8 signal sequence (pTRAP-SS) and pTRAP mixed at a ratio of  $10^{-2}$ ,  $10^{-3}$ ,  $10^{-4}$ ,  $10^{-5}$  or  $10^{-6}$ . As expected, the CD8 positive cells were gradually enriched in the control cell populations transduced even with the lowest percentage ( $10^{-6}$ ) of pTRAP-SS, but not in that transduced with the pTRAP vector only (Table 2 and Fig. 1B). This results indicated that the selection was successful.

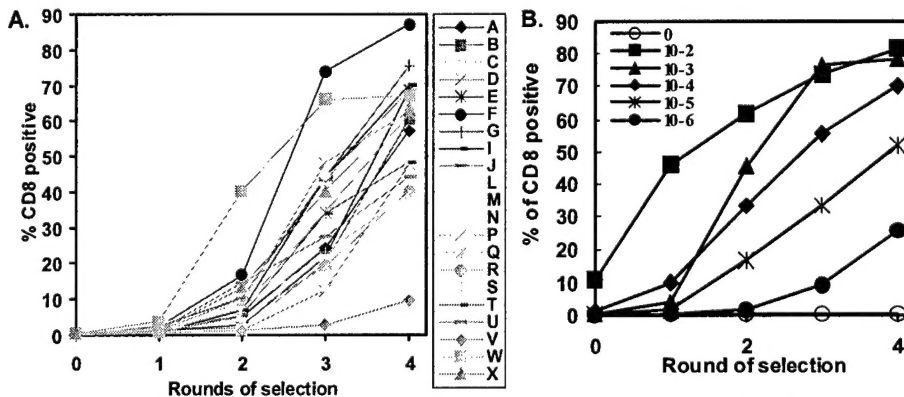


| POOL # | Pre-Selection | First Round | Second Round | Third Round | Fourth Round |
|--------|---------------|-------------|--------------|-------------|--------------|
| A      | 0.2           | 1.8         | 5.2          | 24.1        | 57.3         |
| B      | 0.1           | 1.2         | 2            | 20          | 60.7         |
| C      | 0.1           | 1.5         | 6.7          | 43.3        | 68           |
| D      | 0.1           | 0.9         | 1.1          | 12.3        | 47.5         |
| E      | 0.1           | 1.1         | 2.9          | 21.9        | 68.6         |
| F      | 0.2           | 2.6         | 16.7         | 73.8        | 87.2         |
| G      | 0.6           | 2.5         | 10.4         | 44.3        | 75.6         |
| I      | 0.1           | 0.6         | 6.8          | 44.3        | 70.2         |
| J      | 0.2           | 1           | 10.2         | 43.3        | 68.6         |
| L      | 0.2           | 2.7         | 12.3         | 44.6        | 63.5         |
| M      | 0.2           | 0.9         | 1.5          | 21          | 49.2         |
| N      | 0.1           | 0.8         | 1.4          | 13.5        | 49.8         |
| P      | 0.2           | 2           | 9            | 48.3        | 62.5         |
| Q      | 0.2           | 0.8         | 7.2          | 34.3        | 61.8         |
| R      | 0.2           | 1.2         | 4.6          | 19.1        | 40.1         |
| S      | 0.1           | 0.7         | 6.1          | 23.1        | 39           |
| T      | 0.1           | 0.5         | 5.5          | 34          | 48.4         |
| U      | 0.1           | 1.1         | 14.5         | 27.5        | 44.2         |
| V      | 0             | 0.6         | 1.3          | 2.7         | 9.5          |
| W      | 0.3           | 3.8         | 40.1         | 66.2*       | 67.1         |
| X      | 0.2           | 0.9         | 13.2         | 40.3        | 63.7         |

**Table 1.** % of CD8 positive cells after each round of selection with magnetic beads coupled to an anti-CD8 antibody. The initial cell populations were transduced with pools of the secretion trap library from human breast tumors. After each round of selection, a portion of the cells was stained with an anti-CD8 antibody coupled to FITC and analyzed by flow cytometry.

|           | Pre-Selection | First Round | Second Round | Third Round | Fourth Round |
|-----------|---------------|-------------|--------------|-------------|--------------|
| 0         | 0.1           | 0.4         | 0.3          | 0.4         | 0.5          |
| $10^{-2}$ | 10.6          | 46.1        | 62           | 73.6        | 81.5         |
| $10^{-3}$ | 0.3           | 4.1         | 45.4         | 76.5        | 78.5         |
| $10^{-4}$ | 1             | 9.9         | 33.4         | 55.4        | 70.2         |
| $10^{-5}$ | 0.4           | 1.4         | 16.5         | 33.2        | 51.8         |
| $10^{-6}$ | 0.1           | 0.3         | 1.4          | 9.2         | 25.8         |

**Table 2.** % of CD8 positive cells after each round of selection with magnetic beads coupled to an anti-CD8 antibody. The initial cell populations were transduced with a mixture of the pTRAP vector containing a CD8 signal sequence and the pTRAP vector mixed at a ratio of 0,  $10^{-2}$ ,  $10^{-3}$ ,  $10^{-4}$ ,  $10^{-5}$  and  $10^{-6}$ . After selection, part of the cells was stained with an anti-CD8 antibody coupled to FITC and analyzed by flow cytometry.



**Fig. 2** Selection of CD8 positive cells from 293 cells transduced with a secretion trap library derived from human breast tumors or control plasmids. % of CD8 positive cells after each round of selection with magnetic beads coupled to an anti-CD8 antibody. Cells were transduced with pools of a secretion trap library from human breast tumors (A) or with the pTRAP vector containing a CD8 signal sequence and the pTRAP vector mixed at a ratio of 0,  $10^{-2}$ ,  $10^{-3}$ ,  $10^{-4}$ ,  $10^{-5}$  and  $10^{-6}$ .



### 3. Analysis of the trafficked protein library recovered from the secretion-trap screen.

Provirus were recovered from the 19 pools (A-G, I, J, L-N, P-R, T, U, W and X) in which CD8 positive cells had been enriched to more than 40%. Restriction analysis of randomly selected clones from each pool indicated that most of these pools were represented by multiple cDNA species of different sizes, with the exception of 4 pools that seemed to be dominated by 2-3 different types of cDNA. Therefore, it seems that our new strategy has solved the problem we encountered in our first screen, and that the domination of a small number of clones did not occur in the majority of the pools in the second screen.

#### Known trafficked proteins with signal peptides (48)

Ig heavy chain  
MHC class II lymphocyte antigen  
Apo D  
Serglycin  
Amphiglycan  
CD8  
RAGE (receptor for advanced glycation end product)  
5E10 Antigen - Thy-1  
26kD cell surface protein TAPA-1  
CD44  
Germline Ig lambda light chain variable region  
chromogranin A  
Cathepsin D  
TRAP (translocon-associated protein)  
CD14  
Chondroadherin  
E-cadherin  
Wnt-13  
Pro-1(I) collagen chain (Col1A1)  
Mucin  
MGC24  
SPARC/Osteonectin/BM-40  
Biglycan (hPGI)  
Microfibril associated glycoprotein (MFAP2)  
PDGF-A  
Membrane cofactor protein (MCP, CD46)  
Matrix Gla protein  
Phospholipid transfer protein (PLTP)  
Preprothrotropin-releasing hormone  
BiP protein  
90k tumor associated antigen  
Basigin  
Carbonic anhydrase precursor (CA 12)  
Vacuolar H<sup>+</sup> ATPase proton channel subunit

Chemokine HCC-1  
Aspartyl beta-hydroxylase  
Immunoglobulin J-chain  
Semaphorin 2C  
CD24  
lipophilin B  
Sulphoglucosamine sulphohydrolase (NSS)

Pancreatic cancer specific XS147  
Estrogen regulated LIV-1 from MCF-7 cells  
Human testis enhanced gene transcript protein  
Lysosomal pepstatin insensitive protease (CLN2)  
Sushi repeat containing protein x-chromosome  
Oligosaccharyltransferase  
Thyroid hormone binding protein (p55)

#### Homologs of known trafficked proteins (4)

Rat salivary protein 1  
Circetulus griseus H411 precursor  
Rabbit Angiotensin-converting enzyme (ACE-P)  
Pig uteroferrin

#### Known trafficked proteins without signal peptides (4)

ECE (Endothelin-converting enzyme) (type II)  
Prothymosin  
Glucose transporter (HepG2)  
Thromboxane A-2 receptor

#### Novel proteins with signal peptides (38)

#### Novel trafficked proteins without signal peptides (5)

**Table 3.** A list of trafficked proteins found in the trafficked protein library recovered from our screen.

We randomly selected 20 clones from each of the 15 pools that were represented by multiple cDNA species, as revealed by restriction analysis. Plasmids were isolated from these clones by miniprep and subjected to sequencing analysis. Among the 300 clones that were sequenced, 114 did not yield any information due to poor sequencing reactions. These might be due to poor DNA quality or contamination of unrelated vectors. Sequences were obtained from the remaining 186 clones, and were analyzed by the BLAST program to reveal their identity and by the SignalP program to predict the existence of signal peptides. Among these clones, 59 (32%) encoded known proteins that were not trafficked, such as ribosomal proteins, translational initiation factors and others. The remaining 127 (68%) clones represented 99 different cDNA species, among which

48 were known trafficked proteins with signal peptides, 4 were homologues of known trafficked proteins with signal peptides, 4 were known trafficked proteins without a consensus signal peptide, 38 were novel proteins with signal peptides (as predicted by the SignalP program), and 10 were novel proteins without signal peptides (Table 3).

#### 4. Analysis of novel proteins without a consensus signal peptide.

We found 10 novel proteins without signal peptides in the recovered trafficked protein library. These may be type II or III trafficked proteins that lack consensus signal peptides and cleavage sites. Alternatively, they may be intracellular proteins that are present in the library, since the CD8 positive cells were only enriched to 40-90% before the library was recovered. To differentiate these possibilities, we retransduced these recovered clones into 293 cells and tested whether these sequences could localize the CD8 reporter to cell surface. Our results revealed that 5 of these 10 clones rendered the 293 cells CD8 positive, indicating that these 5 sequences could function as signal peptides. The remaining 5 clones failed to traffic the CD8 reporter and probably represented intracellular proteins.

Based on the sequencing analysis of these 186 clones from the trafficked protein library, we conclude that approximately 68% of the library is represented by trafficked proteins. Among the trafficked proteins contained in the library, 52.5% are known

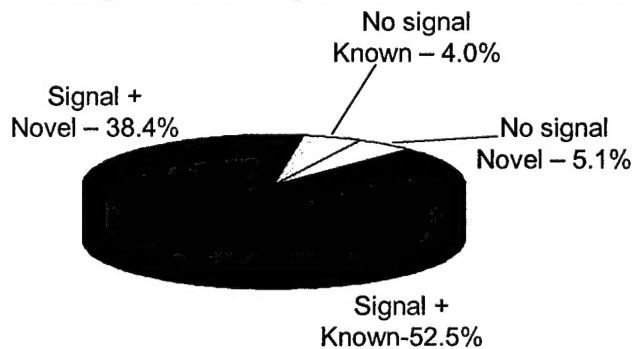


Fig. 3 A breakdown of the trafficked proteins found in the trafficked protein library. Among them, 52.5% are known trafficked proteins with consensus signal peptides, 4.0% are known trafficked proteins without consensus signal peptides, 38.4% are novel proteins with consensus signal peptides, and 5.1% are novel trafficked proteins without signal peptides.

trafficked proteins with consensus signal peptides, 4.0% are known trafficked proteins without consensus signal peptides, 38.4% are novel proteins with consensus signal peptides, and 5.1% are novel trafficked proteins without signal peptides (Fig. 3).

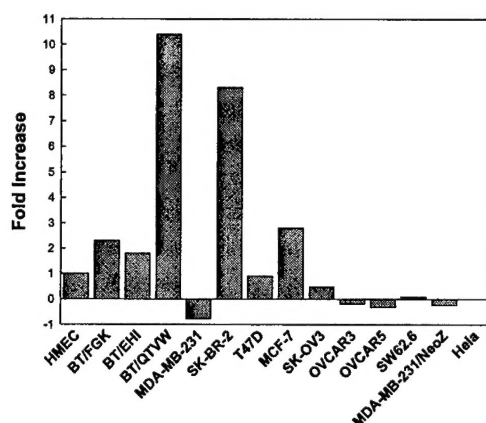
The known trafficked proteins found in our library include receptors, ion channels, transporters, proteases, cytokines,

chemokines, adhesion molecules and carrier proteins.

#### 5. Analysis of LIV-1 expression in human breast tumor cells.

Our ultimate goal is to identify trafficked proteins that are overexpressed in human breast tumors. We noticed that some of the trafficked proteins identified during the analysis of the trafficked protein library had already been reported to be up-regulated in human breast carcinomas. For example, the increased expression of LIV-1 is associated with estrogen-positive human breast cancer (1, 2).

To confirm that our trafficked protein library contained genes that were overexpressed in human breast tumors, we compared the LIV-1 mRNA levels in human breast tumor tissues and cell lines with that in HMEC, a normal human mammary



**Fig. 4** LIV-1 is overexpressed in human breast carcinoma tissues and certain human breast tumor cell lines. 10  $\mu$ g of RNA isolated from normal human mammary epithelial cells (HMEC), 3 mixtures of human breast tumors samples (FGK, TEH and QTVW), 4 human breast carcinoma cell lines MDA-MB-231, SK-BR-2, T47D, and MCF-7, 4 human ovarian tumor cell lines SK-OV3, OVCAR3, OVCAR5 and SW62.6, and a human cervical carcinoma cell line Hela were separated by electrophoresis, transferred to a HybondN+ membrane and hybridized with a  $^{32}$ P-labeled LIV-1 DNA probe. Signals were quantified by Phosphorimager, and the fold increase of LIV-1 signals over that in HMEC was calculated after normalization to the GAPDH signals.

epithelial cell line, by Northern blot analysis (Fig. 4). LIV-1 mRNA was up-regulated in all 3 RNA samples isolated from mixtures of human breast tumor tissues, and 2 (SK-BR-2 and MCF-7) of the 4 human breast carcinoma cell lines by 2- to 10-fold. LIV-1 expression was not increased in Hela human cervical carcinoma cells and 4 human ovarian tumor cell lines (SK-OV3, OVCAR3, OVCAR5 and SW62.6) that we examined. Therefore, increased LIV-1 expression seemed to be specific to human breast carcinomas. This result

indicated that our trafficked protein library includes genes that are overexpressed in human breast tumors, and thus is a good source for the identification of diagnostic markers for human breast cancer. We are currently performing microarray analysis to identify those trafficked proteins with increased expression in human breast carcinomas.

### Key research accomplishments

We have obtained a library of trafficked proteins from human breast tumors in which the signal sequences have been enriched to at least 60%.

### Reportable outcomes

We have obtained a library of trafficked proteins from human breast tumors in which the signal sequences have been enriched to at least 60%.

### Conclusions

We have obtained a human breast tumor library in which the signal sequences have been enriched. Based on the sequencing analysis of a limited number of clones from the library, we conclude that approximately 68% of the library is represented by trafficked proteins. Among the trafficked proteins contained in the library, 52.5% are known trafficked proteins with consensus signal peptides, 4.0% are known trafficked

proteins without consensus signal peptides, 38.4% are novel proteins with consensus signal peptides, and 5.1% are novel trafficked proteins without signal peptides. The known trafficked proteins found in our library include receptors, ion channels, transporters, proteases, cytokines, chemokines, adhesion molecules and carrier proteins. A preliminary analysis revealed that this library contained trafficked proteins with increased expression in human breast carcinomas. A micro-array analysis of the trafficked protein library is currently being carried out to systematically identify trafficked proteins that are overexpressed in human breast tumors.

## References

1. Manning, D. L., Daly, R. J., Lord, P. G., Kelly, K. F. and Green, C. D. (1988) Effects of oestrogen on the expression of a 4.4 kb mRNA in the ZR-75-1 human breast cancer cell line. *Mol. Cell. Endocrinol.* 59, 205-212
2. Taylor, K.M., Morgan, H.E., Johnson, A., Hadley, L.J. and Nicholson, R.I. (2003) Structure-function analysis of LIV-1, the breast cancer-associated protein that belongs to a new subfamily of zinc transporters. *Biochem J.* 375:51-9.

## Appendices

### Personnel

|                            |                        |
|----------------------------|------------------------|
| Peiqing Sun, Ph.D.         | Principal Investigator |
| Rong Liao                  | Research Technician    |
| Bettina Moser Ph.D.        | Research Associate     |
| Christian Kannemeier Ph.D. | Research Associate     |

### Publications

1. Wang WJ, Chen X, Liao R, Deng Q, Zhou JJ, Huang Sand Sun P, 2002. Sequential Activation of the MEK-Extracellular Signal-Regulated Kinase and MKK3/6-p38 Mitogen-Activated Protein Kinase Pathways Mediates Oncogenic *ras*-Induced Premature Senescence. *Mol. Cell Biol.* 22:3389-3403.
2. Li X, Zhao Q, Liao R, Sun P, Wu X, 2003. The SCF(Skp2) ubiquitin ligase complex interacts with the human replication licensing factor Cdt1 and regulates Cdt1 degradation. *J Biol Chem.* 278:30854-8
3. Deng Q, Liao R, Wu BL and Sun P, 2004. High-intensity ras signaling induces premature senescence by activating p38 pathway in primary human fibroblasts. *J. Biol. Chem.* 279:1050-9 (Epub 2003 Oct. 29).
4. Bian D, Su S, Mahanivong C, Chen RK, Han Q, Pan ZK, Sun P and Huang H, 2004. Lysophosphatidic Acid Stimulates Ovarian Cancer Cell Migration via a Ras-MEK Kinase 1 Pathway. *Cancer Res.* 64:4209-17.
5. de Parseval A, Ngo S, Sun P and Elder JH, 2004. Factors that increase the effective concentration of CXCR4 dictate feline immunodeficiency virus tropism and kinetics of replication. *J Virol.* 78:9132-43.
6. de Parseval A, Chatterji U, Sun P and Elder JH, 2004. Feline immunodeficiency virus targets activated CD4+ T cells by using CD134 as a binding receptor. *Proc. Natl. Acad. Sci.* 101:13044-9.

### Meeting abstracts

Era of Hope Department of Defense Breast Cancer Research Program Meeting, Orlando FL, Sept. 25-28, 2002

### IDENTIFICATION OF POTENTIAL BIOMARKERS FOR THE EARLY DIAGNOSIS OF BREAST CANCER

Peiqing Sun, The Scripps Research Institute

Breast cancer is one of the leading causes of cancer mortality in woman. The successful treatment of the disease depends on early diagnosis. Currently the diagnostic tools for breast cancer are limited, and often suffer from drawbacks. For example, mammography can miss small lesions, and sometimes can induce tumors in certain patients. Tests of biopsied tumor samples can only be performed in symptomatic patients when the tumor

has been identified. Certain blood-borne markers have been used in affordable, routine serum tests to diagnose cancer in asymptomatic patients. However, these markers, when used individually, can detect the presence of only a certain percentage, but not all, of a particular type of tumors. This argues that a panel of blood-borne markers should be identified and used for the diagnosis of each type of cancer. It is conceivable that multiple blood-borne markers, when used in combination, could predict the existence of tumors more accurately and more successfully. Once the presence of a tumor is determined, cell surface markers can be used in imaging analysis to identify the location of the lesion.

We propose to systematically isolate secreted and cell surface proteins (trafficked proteins) with increased expression in early stage breast tumors. These proteins are candidates for blood-borne markers and cell surface markers that can be used in the routine screening of early stage breast cancer. We will first isolate all secreted and cell surface proteins from breast tumors of multiple patients, using a functional approach we have designed and validated. Next, the expression levels of these proteins in normal and early stage breast tumor tissues will be compared, and those with increased expression in tumors will be identified and analyzed. These studies will generate a pool of potential biomarkers, which can be evaluated through further studies for their use in the early diagnosis of breast cancer.